

## Cuantificación de las citas textuales del *Diccionario de Autoridades*\*

Guillermo Rojo  
Universidad de Santiago de Compostela

### 1. Introducción

Gracias al trabajo realizado en estos últimos años por diversos autores (cf., entre muchos otros, Lázaro Carreter 1972, Ruhstaller 2000, Freixas 2003 y 2010, Álvarez de Miranda 2005), conocemos ya bastante bien los aspectos fundamentales relacionados con la génesis y el desarrollo del conocido como *Diccionario de Autoridades* (DA). Sin embargo, estamos todavía bastante lejos de tener una visión suficientemente detallada de su contenido. Esa discrepancia se manifiesta nítidamente en el contraste entre la relevancia concedida al uso de ejemplos reales, ilustrativos del uso de las palabras, como elemento configurador del DA y lo poco que sabemos acerca de los autores y obras utilizados en la ejemplificación y la frecuencia con que se emplean textos de las diferentes fuentes.

Tal aparente falta de interés es consecuencia directa de las dificultades existentes para acometer el análisis de las fuentes utilizadas y su grado de empleo en una obra tan amplia como el DA. Ha habido trabajos centrados en algunos autores, como los de Jammes (1996) para las citas de Góngora, Bravo (1998) con las de Villegas, Florit (2004) para las de Tirso de Molina, Candelas (2004) con respecto a las de Quevedo en el primer tomo del DA o Álvarez de Miranda (2004) con Quevedo y Cervantes en una muestra de 200 páginas de cada uno de los tomos. Otros han estudiado el grado de utilización de algunas obras, como ha hecho Prieto García-Seco (2015) con *La pícaro Justina*. También ha habido análisis de cierto tipo de léxico, como han hecho Gutiérrez Rodilla (1994) con el relacionado con la medicina, Salvador Rosa (1985) con las palabras que llevan marca geográfica, Aliaga (1994) para las características de Aragón, Campos y Pérez Pascual (2012) para las relacionadas con Galicia, Ruhstaller (2004 y 2006), Prieto García-Seco (2010) para las 'jocosas e inventadas' y Breuer (2007) ha localizado todas las que llevan una marca. Por fin, Freixas (2003, 2010: 267 y sigs.) ha trabajado con la totalidad de las citas y referencias de dos muestras distintas, que suponen, en conjunto, alrededor del 5% de la obra. El resultado de su trabajo, llevado a cabo de forma manual, traza un panorama bastante completo del grado en que los académicos fundadores emplearon los textos seleccionados y a cuáles recurrieron con más frecuencia, pero deja ver también que la parte de la obra no incluida en las muestras puede contener

---

\* Texto construido sobre los apartados generales de Rojo (2014 y 2016). © G. Rojo. Marzo de 2016.

elementos no presentes en ellas y, por supuesto, presentar rasgos cuantitativos diferentes.

Evidentemente, la solución no pasa por invertir más tiempo en el muy tedioso análisis manual de la obra. El estudio completo del DA solo puede plantearse mediante la utilización de una versión digitalizada que permita automatizar al menos la parte más mecánica de esta tarea. No será necesario insistir en que la digitalización del DA es, en sí misma, un trabajo muy lento y costoso, puesto que resulta inviable la utilización de programas de reconocimiento de caracteres y, por tanto, tiene que ser llevada a cabo introduciendo a mano el texto de la obra. En paralelo a la construcción del *Corpus diacrónico del español* (CORDE) y pensando sobre todo en la preparación de materiales para el *Nuevo diccionario histórico del español* (NDHE), la Real Academia Española acometió la digitalización manual de los seis tomos del DA. El resultado está ya a disposición pública en la página electrónica de la RAE (<http://web.frl.es/DA.html>), con una aplicación de búsqueda, desarrollada en el Instituto de investigación Rafael Lapesa de la RAE, que permite recuperar la información correspondiente al texto de las entradas.<sup>1</sup>

Esta versión digital del DA está todavía muy poco codificada, con lo que las posibilidades de recuperación de información son completas solo si se conoce la presentación exacta de aquello que se quiere obtener. Dado que he tenido la posibilidad de manejar directamente el texto digitalizado, he aplicado algunas técnicas básicas de recuperación de información textual contenida en las entradas a la identificación de los autores y las obras citados en el texto y, posteriormente, al recuento de las apariciones de cada uno de ellos. Espero, de este modo, facilitar la parte más pesada del trabajo de recopilación de datos a quienes estén interesados en el análisis de la presencia en el DA de ciertos autores en general o de sus diferentes obras.

## 2. El procedimiento

Los resultados obtenidos mediante análisis manual del contenido del DA son muy fiables, puesto que utilizan el conocimiento técnico de la persona que ha llevado a cabo los recuentos, así como la capacidad de abstracción necesaria para reunir en una entidad única referencias que difieren ligeramente, bien sea por inconsistencia en las abreviaturas utilizadas bien por erratas en la impresión. Sin embargo, la proyección que se puede obtener de muestras de ese tipo tiene, por su propio carácter, los riesgos inherentes a esta técnica, que no puede prever las peculiaridades de cada parte de la obra<sup>2</sup> y, por consiguiente, corre el riesgo de producir proyecciones erróneas en la

---

1 Es, pues, algo muy diferente de lo que se puede conseguir con el *Nuevo tesoro lexicográfico*. En esa otra aplicación, utilísima, la consulta se hace por los lemas y la respuesta devuelve la imagen de la entrada correspondiente.

2 Aspecto importante, puesto que, como es bien sabido, la responsabilidad de la recogida de papeletas y la redacción de las entradas estuvo a cargo de diferentes académicos, con lo que ello puede suponer de discrepancia en la

generalización de los cálculos y en la falta de detección de autores y obras citados en entradas situadas en zonas de la obra que no forman parte de las muestras analizadas. Por el contrario, los procedimientos automáticos de recuperación y cuantificación de información permiten trabajar con comodidad sobre textos de gran volumen y son capaces de localizar todas las referencias que tienen una determinada configuración formal, pero tienen otros inconvenientes. En primer lugar, exigen la digitalización cuidadosa del texto, proceso complejo y costoso, puesto que debe hacerse manualmente, ya que la utilización de programas de reconocimiento óptico de caracteres no da, por el momento, buenos resultados en textos impresos con la tipografía del DA. Por otro lado, la recuperación automática de la información pertinente tiene que enfrentarse con la posible incongruencia de los datos necesarios, que no siempre aparecen del mismo modo, con las mismas convenciones, ni pueden solucionar las discrepancias debidas a erratas de imprenta o fallos en el proceso de digitalización, etc. La utilización de esta vía tiene, por tanto, el riesgo de no detectar citas que existen en el texto y también de hacer mal los recuentos, por exceso o por defecto.<sup>3</sup>

Una parte considerable de estos problemas desaparece cuando el texto digitalizado es revisado y la información contenida en las entradas es estructurada y codificada al modo en que se hace en los diccionarios electrónicos. Ese proceso, todavía más largo, más complejo y más caro que la digitalización inicial, no ha sido acometida todavía para el DA, de modo que la detección de autores y obras por un lado y los recuentos automáticos tienen que enfrentarse con bastantes problemas que arrojan una cierta inseguridad sobre los resultados obtenidos. A pesar de ello, creo que lo que se puede hacer ya en la situación actual arroja nueva luz sobre las características y el valor de la obra y permite llegar a una visión más amplia de lo que supone en la historia de la lexicografía, no solo la española.

Como es bien sabido, los académicos parten de una lista inicial de autores y textos que consideran los más representativos y adecuados para ilustrar el uso de las palabras en sus diferentes acepciones. A esa “[l]ista de los autores elegidos por la Real Academia Española, para el uso de las voces y modos de hablar” (DA: I, lxxxv) se añaden muchos otros que la complementan y todos ellos figuran en la “[e]xplicación de las abreviaturas de los nombres de Autores y obras que van citados en este [...] Tomo” (DA: I, lxxxix) que aparecen al comienzo de cada uno de los seis volúmenes. Por tanto, el diseño de la recogida de información para hacer el recuento parece

---

configuración de las entradas, número de citas utilizadas, etc.

3 En la recuperación automática de información se utilizan habitualmente dos medidas. La precisión (ingl. *precision*) refleja la relación entre el número de aciertos y el número de elementos recuperados. La sensibilidad (ingl. *recall*) se refiere a la relación entre el número de elementos recuperados correctamente y el número de elementos de esa clase existentes en los documentos sobre los que se hace la búsqueda. En este caso, está claro que la precisión es muy alta y los problemas pueden estar en la sensibilidad.

sencillo. En principio, sería suficiente con producir una lista única de todas las abreviaturas que figuran en las seis relaciones y preparar luego una rutina capaz de revisar el texto del DA haciendo el recuento de las veces que es mencionado cada autor (o cada pareja autor-obra si se pretende llegar hasta ese grado de detalle). No es complicado porque los académicos diseñaron un sistema de abreviaturas realmente claro: responden básicamente al tipo QUEV. Alguac./.<sup>4</sup> Esto es, la abreviatura elegida para el autor con versal inicial y versalitas para el resto de la clave, un punto indicador de la abreviatura, un espacio en blanco, y clave de la obra con versal inicial y resto en minúscula, también con puntos para cada palabra abreviada, y todo ello en redonda. En la edición digitalizada que he podido manejar, la diferencia entre versal y versalita ha sido eliminada y son todas versales. A la abreviatura de la obra puede seguir la indicación del lugar (página, folio, capítulo, etc.) del que se ha obtenido la cita y, finalmente, el texto del ejemplo, con la palabra correspondiente al lema destacada tipográficamente. Como es natural, tanto la clave de autor como la de obra pueden constar de varias palabras, pero las convenciones se mantienen (FR. L. DE GRAN. Trat. de la Orac.), de modo que es igualmente sencillo localizarlas. Por tanto, restringiéndonos ya a los recuentos de autores, la detección automática de las entradas en las que se contiene, por ejemplo, la cadena QUEV. no es tarea complicada y, además, gracias al empleo de mayúsculas y el punto de la abreviatura no puede ser confundida con otras palabras que contengan esa secuencia de caracteres en su interior. Algo muy parecido ocurriría en caso de que se quisiera hacer el recuento de autor y obra.

Sobre esa convención general, realmente bien concebida, como he dicho, se añaden otras convenciones parciales, que enumero rápidamente, sobre todo por si pueden ser de ayuda para quienes deseen localizar autores u obras en el texto del DA que se puede consultar en la página de la RAE.<sup>5</sup>

- (a) Puede faltar la indicación de obra cuando, por alguna razón, se considera innecesario hacerlo, sea porque, a juicio de los académicos, no caben dudas acerca de cuál es la obra (así COVARR. para el *Thesoro de la lengua castellana*), sea porque no siempre se diferencia en el interior del conjunto de las obras de ese autor (AMBR. MOR./, por ejemplo).<sup>6</sup>

---

4 Dado que es necesario mantener los detalles tipográficos tal como aparecen en el texto impreso o bien en la edición digitalizada, para referirme a ellas en el cuerpo del texto utilizo una familia distinta, de espacio fijo. Puesto que hay puntos y comas que forman parte de la clave, utilizo una barra (/) para separar las distintas claves y también su final en los casos necesarios.

5 Como he indicado, el texto impreso del DA distingue, casi siempre de forma cuidadosa, entre versales y versalitas, pero la versión digitalizada ha prescindido de esta diferencia. Por tanto, a partir de este punto utilizo simplemente la diferencia entre mayúsculas y minúsculas en la referencia a las claves de autor y obra.

6 Cf. Rojo (2016, apdo. 2) para el caso de este autor. En muchos otros casos, el desarrollo de las abreviaturas que figuran en las listas dice “sus obras”, “sus comedias”, “su obra poética”, etc. sin más detalles, pero el texto de las

- (b) La serie de caracteres en mayúscula (versal inicial y versalita en la edición impresa) puede corresponder no a un autor, sino a una obra en el caso de las anónimas (Boc. DE OR.), textos de carácter especial (ARANC. para *Los nuevos aranceles*) o bien cuando los redactores consideran más claro o más adecuado aludir directamente a la obra (C. LUCAN. para el *Conde Lucanor* o bien LAZAR. DE TORM./, a pesar de que lo atribuyen a Diego Hurtado de Mendoza).
- (c) En principio, el punto es indicativo de abreviatura, de modo que es sustituido por una coma cuando se utiliza el nombre completo (GARAI, Cart. para “Blasco de Garai: Cartas en refranes”, que figura en el tomo III).

La obtención de estadísticas generales acerca de los autores, las obras y la frecuencia con que aparecen en la totalidad del DA, inviable por procedimientos manuales, se convierte en algo relativamente sencillo desde el momento en que existe la posibilidad de procesar informáticamente la lista de claves de autores (y obras si se considera oportuno) utilizadas en el DA y el texto de las entradas. En definitiva, se trata simplemente de saber cuántas veces aparece la clave de cada uno de los autores (o autores y obras) de la lista en el texto de cada uno de los seis tomos. En consecuencia, el primer paso consistió en fundir en un documento único las seis relaciones de autores (y obras mencionadas directamente), eliminando las claves de obra y el desarrollo de autor y obra. De ahí salió una lista formada por 479 elementos, en la que, como es lógico, figuran también las variantes incluidas en las seis relaciones (por ejemplo, AYALA,/ o BOCAD. DE OR./ en el tomo I, frente a AYAL./ y BOC. DE OR./ en el tomo II), así como todas las variantes gráficas resultantes de descuidos o erratas (del tipo AVIL./ y Avil./).

Trabajar con una lista única de autores resultante de la fusión de lo que se encuentra en las seis relaciones iniciales permite obtener recuentos fiables incluso en aquellos casos en los que la congruencia entre claves y entradas es bastante reducida. Por ejemplo, en la lista del tomo I, Teodoro Ardemans aparece abreviado como ARDEMANS,/. A pesar de ello, no hay ni una sola entrada que contenga esta clave. La única cita de Ardemans que se encuentra en el tomo I puede ser recuperada automáticamente gracias a que en las listas de otros tomos este autor figura con ARDEM./.

Pero no todos los casos son de este tipo. Como es bien sabido, la práctica de la papeletización de los textos, redacción y revisión de las entradas introdujo en el texto del DA una gran cantidad de claves que no están en ninguna de las listas iniciales. Por ejemplo, las obras de Basilio Barén de Soto figuran en las listas siempre, de acuerdo con el diseño general, como

---

entradas indica casi siempre la obra concreta de la que procede la cita. Naturalmente, esa discrepancia complica la recuperación de la información relativa a las obras.

BAREN,/ (que aparece en 244 citas). Sin embargo, hay otros 82 fragmentos de este autor con la clave BAREN./ y resulta especialmente llamativo que la distribución de las que aparecen en el tomo I sea 49 para la variante con punto (la contraria al sistema, que no está en la lista de abreviaturas) y 1 para la variante con coma (la que figura en la relación). Se mantiene el sistema general, pero se cambian las abreviaturas en las que aparecen en la lista para el *Doctrinal de caballeros*: DOCTR. DE CAB./ DOCTR. DE CABALL./ DOCTRIN. DE CAB./ DOCTRIN. DE CABALLER./ . En diferentes entradas del DA aparecen, además, DOCT. DE CAB./ DOCT. DE CABALL./ DOCTR DE CAB./ DOCTR. DE CABALLER./ .

El análisis del texto del DA muestra, por otro lado, la existencia de un grupo importante de autores y obras que no figuran en la lista de abreviaturas. A algunos de ellos alude Álvarez de Miranda (2005: 27) en su revisión de las entradas de la letra B, pero el contraste de la lista general de abreviaturas con las se incluyen aquí mostrará muchos otros casos. Como es lógico, esas ausencias no se dan en autores (y obras mencionadas directamente) del grupo de los más citados, pero todos juntos pueden suponer una cantidad considerable de citas y, sobre todo, no detectarlos supone la pérdida de una información que puede ser muy valiosa. Así, la revisión de la zona correspondientes a las *Ordinaciones y Ordenamientos* del apéndice I de Rojo (2014) muestra, entre otras, las claves ORDEN. DE ALCAÑ./ ORDEN. DE ALCAÑIZ./ ORDEN. DE DAROC./ ORDEN DE DAROCA,/ ORDEN. DE HUES./ ORDEN. DE HUESC./ ORDEN. DE LA CHANC. DE GRAN./ ORDEN DE LA INFANT./ ORDEN. DE LA PLAN. MAYOR DEL EXERC./ ORDEN. DEL EXERC. DE FLAND./ ORDEN DEL REGIM. DE GUARD./ ORDEN DEL VALLE DE ARAM./ ORDEN. PARA EL DIRECT. DE LA INFANT./ ORDEN. PARA EL EXERC. DE FLAND./ ORDEN. PARA LA FORMAC. DE LAS MILIC. DEL REINO,/, que no aparecen registradas en la lista de abreviaturas ni en esta forma ni en ninguna variante.

Hay que tener en cuenta, además, las diferencias derivadas de las erratas que se produjeron en el proceso de impresión. Incómodas, pero sin repercusiones graves para las lecturas humanas, constituyen un problema más grave para las lecturas automáticas en la medida en que hayan pasado al texto digitalizado. Este último añade, por su parte, una cierta cantidad de errores resultado de la digitación manual<sup>7</sup>. Además, en, al menos, algunos de los tomos se introdujeron marcas que, dada la insuficiente codificación del texto, estorban en la recuperación de los candidatos a claves de autor y obra. Por ejemplo, en algunos de los tomos (pero no en todos) se introdujo la marca [w] al inicio de las referencias a un autor, tanto en las citas en sentido estricto como en las referencias indirectas (las

---

<sup>7</sup> Algo perfectamente esperable por el volumen de la obra. Según mis recuentos, el texto de las entradas de los seis tomos (esto es, sin preliminares) contiene 179 860 párrafos, 3 975 784 palabras y 23 701 536 caracteres.

que Freixas (2003, 2010) denomina 'fuentes' (cf. Rojo 2014: 147 y sigs.). Salvo casos muy especiales, esas marcas no estorban en la recuperación automática de los autores, pero sí lo hace, en cambio, una marca destinada a indicar el cambio de página en medio de una entrada, del tipo CALD. [v.546] Aut. El Pintór de su deshonra,/.<sup>8</sup> Como es evidente, las búsquedas basadas en la expresión CALD. Aut./ fracasarán en casos de ese tipo. En algunos volúmenes se ha alterado también la grafía de las abreviaturas de los autores añadiendo tildes a las vocales, frente a lo que aparece en el texto original.<sup>9</sup> Así, la clave correspondiente a Fray Luis de León aparece como FR. L. DE LEÓN./ FR. L. DE LEÓN,/ casi 400 veces en el texto digitalizado, frente a cero en el texto impreso (cf. Rojo 2014: apéndice I). Evidentemente, la diferencia entre las grafías con tilde y sin ella puede influir en la recuperación de resultados.<sup>10</sup>

En consecuencia, aplicando lo ya expuesto acerca del sistema empleado habitualmente por los académicos para construir las referencias, preparé una rutina que extrajo del texto del DA todas las secuencias que podrían corresponder a claves de autores y obras mencionadas directamente. Después de la eliminación de repeticiones, revisé atentamente esa lista para eliminar de ella todas las que se ajustaban formalmente a los criterios utilizados, pero no eran realmente claves de autor. La depuración y posterior unión de las dos listas produjo una relación de 988 elementos. Dado que la lista de abreviaturas extraída de las relaciones iniciales tenía 479 claves, esta rutina proporcionó 509 términos de búsqueda nuevos, lo cual da una idea de la importancia cuantitativa y quizá cualitativa de la operación realizada. En otras palabras, en las listas iniciales figura únicamente el 48,48% de las claves realmente empleadas en la obra. Por supuesto, no podemos olvidar que la mayoría de las entradas adicionales son puras variantes tipográficas, erratas de impresión o errores de digitalización, pero es necesario tenerlas en cuenta, puesto que es el único modo de extraer información completa sobre los autores citados en el DA del texto digitalizado que está en la página de la RAE.

El paso siguiente consistió en preparar una segunda rutina, destinada a obtener el número de apariciones de las claves identificadas en el texto de cada uno de los tomos del DA. La

---

8 La primera parte de la definición mencionada contiene, además, un error que no existe en el texto impreso:

*YUNQUE. s. m. Lo mismo que Ayunque. QUEV. Entrem. Eres dixo el Fiscal con esse yunque, y esse martillo, y essos calvos. CALD. [v.546] Aut. El Pintór de su deshonra, en la Loa.*

9 Como es bien sabido, los académicos no dispusieron de mayúsculas con tildes (cf. Lázaro Carreter 1972: 99-100), de modo que para señalar la acentuación de las palabras tuvieron que recurrir a diferentes procedimientos, alguno de los cuales incide sobre la recuperación de la palabra que constituye la cabeza de la entrada (cf. Rojo 2014: 145, nota 14).

10 Eso no sucede en la aplicación de recuperación de datos que funciona en la página de la RAE, que prescinde por completo de la diferencia entre mayúsculas y minúsculas y vocales con tilde y sin ella. Esa solución tiene, como es lógico, ventajas e inconvenientes.

consideración diferenciada de los tomos puede resultar de interés para analizar el peso diferente de algunos autores a lo largo de los años de redacción de la obra. Los resultados han sido revisados para eliminar de los recuentos el efecto de la coincidencia parcial de algunas secuencias cortas con otras más largas que contienen la primera, lo cual distorsiona al alza los recuentos de la corta. Por ejemplo, la búsqueda de la cadena MEND. en los seis volúmenes arroja un total de 1660 casos. Dado que es posible manejar la diferencia entre los casos en los que la cadena en cuestión constituye una “palabra” completa y aquellos en los que forma parte de una “palabra” más larga (como, en este caso, COMEND./ COMEND. GRIEG./ EL COMEND.), la búsqueda de la abreviatura que nos interesa se reduce a 971 casos. Pero el análisis de lo recuperado muestra que esa cifra incluye no solo los casos de MEND./ como clave de autor, sino también los correspondientes a A. MEND./ B. MEND./ D. MEND./ SALAZ. DE MEND./ . Por tanto, es necesario descontar las frecuencias detectadas para esas otras combinaciones de las correspondientes a MEND./, lo cual deja su frecuencia en 682 casos.<sup>11</sup> La complejidad de estos encadenamientos puede ser mayor, como muestra, por ejemplo, el caso de la clave CHRON./, que exige deducir de su frecuencia todos los casos correspondientes a, entre otros muchos, CHRON. GEN./, del que, a su vez, hay que retirar antes los de CHRON. GEN. DE ESP./ CHRON. GEN. DE Esp./.

En el paso siguiente, se trataba de llegar a conocer la frecuencia de utilización de las obras de los diferentes autores. Las dificultades iniciales no eran diferentes de las planteadas para el recuento de autores (y obras mencionadas directamente), pero ya se podía partir de una relación bien contrastada de autores recuperada directamente de los textos y no solo de las listas de abreviaturas. En consecuencia, esta nueva fase consistió en la preparación de una nueva rutina capaz de recorrer el texto de los seis tomos y detectar las cadenas en las que figuraba cada clave de autor (y obra) y extraer también la secuencia de caracteres que viene a continuación, que es la candidata a contener la clave de la obra. Esta rutina produjo una enorme cantidad de líneas (en realidad, tantas como casos detectados) que fueron reducidos automáticamente mediante la supresión de líneas repetidas. A pesar de esa reducción, quedaba todavía una gran cantidad de cadenas que contenían la misma clave de obra, pero también algunas indicaciones adicionales (página, folio, tomo, etc.) o los primeros caracteres del texto de la cita. Fue necesario depurar manualmente esa lista hasta lograr reducirla a las combinaciones de autor y obra utilizadas en el

---

11 Evitar estas agrupaciones en recuentos automáticos supondría una enorme complejidad en la rutina utilizada, de modo que resulta mucho más aconsejable la estrategia de localizar, también automáticamente, los casos potencialmente conflictivos y reajustar manualmente los recuentos. Como es lógico, la aplicación de búsqueda en el texto del DA en la página de la RAE no puede realizar esos ajustes. Véase, por ejemplo, que la búsqueda de la secuencia MEND. devuelve la entrada correspondiente a *calculado, da*, que contiene SALAZ. DE MEND. Como se muestra en Rojo (2016, apdo. 4), este caso tiene complicaciones adicionales.

DA.

El análisis de los resultados obtenidos muestra desajustes similares a los observados en las listas de autores y obras mencionadas directamente, ahora multiplicados como consecuencia del mayor grado de especificidad. En primer lugar, la multiplicación de variantes gráficas diferentes de las que figuran en la lista de abreviaturas. En segundo término, los textos contienen citas de obras de un autor no mencionadas en la lista de abreviaturas y no integrables en conjuntos más amplios (vid. infra). Así, por ejemplo, en la lista de abreviaturas no están los *Entremeses* de Cervantes, que aparecen cuatro veces bajo la clave CERV. Entrem./.

Hay, además, desajustes que surgen como consecuencia de la indicación de autor y obra. Con cierta frecuencia, la lista de abreviaturas contiene indicaciones genéricas del estilo “sus obras”, “su obra poética” o algo por el estilo. En casos de este tipo, la referencia que figura en el texto puede quedarse en la simple indicación del autor, como sucede en las 90 citas de Álvaro Gómez de Ciudad Real (el joven), que son todas ellas del tipo ALV. GOM. Cant. 3. Oct. 15 (con variantes en la clave de autor). Resulta también muy llamativo el caso de Ambrosio de Morales, que es uno de los autores más citados del DA: de los 1045 textos extraídos de sus obras, 886 (es decir, el 84,78%) aparecen con la simple abreviatura del autor, lo cual resulta bastante confuso en una obra tan extensa.<sup>12</sup> Por fin, hay que mencionar casos como el representado por Calderón, que en la lista de abreviaturas aparece como CALD. / CALDER. / seguido de indicaciones genéricas del tipo Aut. / Com. /, pero que en el cuerpo de las entradas lleva casi siempre indicación de la obra concreta de la que procede el ejemplo.<sup>13</sup>

El resultado de todo ello es una gama bastante heterogénea de modos de organización de las obras de los diferentes autores, complicada además por la existencia de variantes gráficas (resultado de las diferentes opciones adoptadas por los académicos, las erratas del texto impreso y los errores de la edición digitalizada). A las diferencias en la mayor o menor amplitud de las obras de un cierto autor se añaden las que resultan de las utilizadas por los académicos para organizarlas. Así, en las listas de abreviaturas, las obras de Cervantes se organizan en *Comedias*, *La Galatea*, *Novelas ejemplares*, el *Persiles*, el *Quijote* y el *Viaje del Parnaso*. Las citas desarrollan las comedias (he registrado cinco distintas, sin tener en cuenta las variantes gráficas) y las novelas (citadas mediante

---

12 En este caso, las referencias genéricas van seguidas normalmente de indicaciones del tipo tom. / o lib. /. Algo diferente es lo que sucede con Pedro de Abarca: 12 de las 23 citas tienen únicamente clave de autor a pesar de que en la lista de abreviaturas figura siempre la referencia a los *Anales de Aragón* con diversas variantes. Además, una de las referencias con indicación de obra dice ABARC. Mahometismo, /, que parece ser el primero de los capítulos de los *Anales*.

13 Por ejemplo, de las 163 citas de CALD. Aut. / CALDER. Aut. /, únicamente 2 carecen de la indicación de la obra concreta.

un número de orden casi siempre), además de añadir los ejemplos tomados de los *Entremeses* (4 citas tomadas de dos textos). El más llamativo es, sin duda, el caso de Calderón. La lista de abreviaturas se limita a diferenciar entre comedias y autores (cf. supra), mientras que las citas llevan casi siempre la indicación de la obra concreta, de modo que se pueden identificar (sin tener en cuenta variantes gráficas) cerca de 60 autos sacramentales y casi 90 comedias diferentes.<sup>14</sup>

El último paso es, por supuesto, pasar la rutina que hace el recuento de casos de cada una de las claves de autor y obra incluidas en la lista por cada uno de los tomos. Se obtiene así una lista paralela a la de autores, pero ahora a un nivel superior de detalle.

### 3. Las listas

En los enlaces que se indican a continuación podrán encontrarse diferentes versiones de las listas de frecuencias de autores y obras mencionadas directamente por un lado y de autores y obras por otro. En cada caso, se incluye una versión en PDF de la lista y también una versión en texto plano y formato *tsv* (campos separados por tabuladores) fácilmente integrable en una hoja de cálculo o una base de datos para facilitar el manejo y reordenación de los datos.

Además, se incluye la lista de abreviaturas resultante de la fusión de las relaciones de los seis tomos. Naturalmente, contiene todas las variantes gráficas, incluidas las que corresponden no solo a la abreviatura, sino a su desarrollo. También en este caso hay una versión en PDF y otra en formato *tsv*.

#### 3.1. Lista unificada de abreviaturas.

Cada uno de los tomos del DA contiene una “[e]xplicación de las abreviaturas de los nombres de Autores y obras que van citados en este [...] tomo”. Como ya se ha indicado, los académicos no siempre respetan la convenciones establecidas por ellos mismos ni mantienen las abreviaturas tal como figuran en la relación. A pesar de ello, no es inútil disponer de una relación en la que, en un listado único, se fundan las aparecidas en las primeras páginas de los seis tomos.

La lista que publico en [http://gramatica.usc.es/~grojo/recursos/dic\\_aut/lista\\_unica\\_abreviaturas.pdf](http://gramatica.usc.es/~grojo/recursos/dic_aut/lista_unica_abreviaturas.pdf) contiene todas las variantes registradas, incluidas aquellas que no se dan en la propia abreviatura, sino en su desarrollo. La misma lista, en texto con formato *tsv*, puede descargarse de

---

<sup>14</sup> Teniendo en cuenta las variantes localizadas, el número de claves correspondientes a Calderón de la Barca se eleva a 429.

[http://gramatica.usc.es/~grojo/recursos/dic\\_aut/lista\\_unica\\_abreviaturas.tsv](http://gramatica.usc.es/~grojo/recursos/dic_aut/lista_unica_abreviaturas.tsv).

### **3.2. Lista de autores y obras mencionadas directamente**

Como se señala en el apartado 2, los redactores del DA no siempre respetan las abreviaturas que incluyen en las relaciones incluidas en cada tomo y, al tiempo, hay una importante cantidad de citas de autores y obras que no figuran en esas listas.

En Rojo (2014, apéndice I) puede encontrarse la primera lista de autores y obras con los datos de frecuencia correspondientes a cada tomo y el total del DA. Desgraciadamente, ese listado contiene algunas inexactitudes producidas por la falta de detección de algunas obras o por errores en el reajuste de frecuencias necesario por la coincidencia parcial de claves. Publico aquí la lista con todas las correcciones que he podido incorporar hasta febrero de 2016. La versión que aparece en [http://gramatica.usc.es/~grojo/recursos/dic\\_aut/recuento\\_autores\\_alf.pdf](http://gramatica.usc.es/~grojo/recursos/dic_aut/recuento_autores_alf.pdf) está ordenada alfabéticamente y la de [http://gramatica.usc.es/~grojo/recursos/dic\\_aut/recuento\\_autores\\_frec.pdf](http://gramatica.usc.es/~grojo/recursos/dic_aut/recuento_autores_frec.pdf) muestra los autores y obras por orden de frecuencia decreciente. Al igual que sucede en la relación publicada en Rojo (2014), las columnas adicionales muestran si la coincidencia parcial de las claves ha obligado a hacer reajustes de frecuencias [R] y si la clave figura en la lista de abreviaturas [LA]. La versión en texto con formato tsv puede descargarse de [http://gramatica.usc.es/~grojo/recursos/dic\\_aut/recuento\\_autores.tsv](http://gramatica.usc.es/~grojo/recursos/dic_aut/recuento_autores.tsv).

### **3.3. Lista de autores y obras**

Mediante el procedimiento descrito en el apartado 2, obtuve la relación de abreviaturas correspondientes a cada una de las parejas autor-obra identificadas previamente y, en un paso posterior, llevé a cabo los recuentos en cada uno de los tomos. El resultado puede verse en [http://gramatica.usc.es/~grojo/recursos/dic\\_aut/recuento\\_autores\\_obras\\_alf.pdf](http://gramatica.usc.es/~grojo/recursos/dic_aut/recuento_autores_obras_alf.pdf) y [http://gramatica.usc.es/~grojo/recursos/dic\\_aut/recuento\\_autores\\_obras\\_frec.pdf](http://gramatica.usc.es/~grojo/recursos/dic_aut/recuento_autores_obras_frec.pdf). Como en la lista anterior, en columnas adicionales se marca si ha habido reajustes en los recuentos [R] y si la clave figura en la lista de abreviaturas [LA]. La versión en texto con formato tsv puede descargarse en [http://gramatica.usc.es/~grojo/recursos/dic\\_aut/recuento\\_autores\\_obras.tsv](http://gramatica.usc.es/~grojo/recursos/dic_aut/recuento_autores_obras.tsv).

## Referencias bibliográficas

- Aliaga, Jiménez, José Luis (1994): *El léxico aragonés en el «Diccionario de Autoridades»*. Zaragoza, Institución «Fernando el Católico».
- Álvarez de Miranda, Pedro (2004): «Quevedo en la lexicografía española», *Edad de Oro*, xxiii, pp. 389-416.
- Álvarez de Miranda, Pedro (2005): «La proeza de un diccionario con textos. Algo más sobre el *Diccionario de Autoridades*», en Martínez Romero, Tomás (ed.): *Les lletres hispaniques als segles xvi, xvii i xviii*. Castellón: Universitat Jaume I / Fundació Germà Colón, pp. 73-92. Cito por su reedición en Álvarez de Miranda (2011), pp. 17-36.
- Álvarez de Miranda, Pedro (2011): *Los diccionarios del español moderno*, Gijón: Trea.
- Álvarez de Miranda (en prensa): «Sobre fuentes no literarias del *Diccionario de Autoridades*, con especial atención a la 'Pragmática de tasas' (1680)». XV Congreso de la Asociación alemana de hispanistas, Heidelberg, 18-22 de marzo de 2015.
- Bravo Vega, Julián (1998): “Estaban M. de Villegas, autoridad léxica”, en García Turza, Claudio, Fabián González Bachiller y Javier Mangado Martínez (eds.): *Actas del IV congreso internacional de historia de la lengua española*. Servicio de Publicaciones de la Univ. de La Rioja, II, 43-52.
- Breuer, Judith (2007): *Normen und normative Konzepte der Real Academia Española. Diasystematische Markierungen im 'Diccionario de Autoridades'*, Bonn: GRIN Verlag.
- Campos Souto, Mar y José Ignacio Pérez Pascual (2010): «Los elementos gallegos en el *Diccionario de Autoridades*». En Nomdedeu Rull, Antoni, Esther Forgas Berdet y Maria Bargalló Escrivà (eds.): *Avances de lexicografía hispánica (I)*, Tarragona: Publicacions URV, pp. 259-270.
- Candelas Colodrón, Manuel Ángel (2004): «Quevedo y el *Diccionario de Autoridades*», en Schwartz, Lía (ed.): *Studies in honor of James O. Crosby*, Newark: Juan de la Cuesta, 69-89.
- DA (Diccionario de Autoridades): *Diccionario de la lengua castellana en que se explica el verdadero sentido de las voces, su naturaleza y calidad, con las frases o modos de hablar, los proverbios o refranes y otras cosas convenientes al uso de la lengua*. 6 tomos. Madrid. 1626-1639. Edición facsimilar, Real Academia Española, 2013.
- Estatutos de la Real Academia Española, en DA, I, xxiii – xxx.
- Florit Durán, Francisco (2001): «La nómina del *Diccionario de Autoridades*: el caso de Tirso de Molina», en Pardo Molina, Irene y Antonio Serrano (eds.): *En torno al teatro del Siglo de Oro. XV jornadas de teatro del Siglo de Oro*, Almería: Instituto de Estudios Almerienses, pp. 71-84.
- Freixas Alás, Margarita (2003): *Las autoridades del primer diccionario de la Real Academia Española*. Tesis doctoral de la UAB. Edición electrónica disponible en <http://www.tdx.cesca.es/TDX-0611104-150443>.
- Freixas Alás, Margarita (2010): *Planta y método del Diccionario de Autoridades. Orígenes de la técnica lexicográfica de la Real Academia Española (1713-1739)*. Anexo 14 de la *Revista de Lexicografía*. Universidade da Coruña.
- Gutiérrez Rodilla, Bertha M. (1994): «Construcción y fuentes utilizadas para los términos médicos en el *Diccionario de Autoridades*», *Revista de Lexicografía*, I, pp. 149-162.
- Jammes, Robert (1996): «Góngora en el *Diccionario de Autoridades*», en *Philologica (Homenaje al profesor Ricardo Senabre)*, Universidad de Extremadura, pp. 247-272.
- Lázaro Carreter, Fernando (1972): *Crónica del Diccionario de Autoridades. (1713-1740)*. Discurso de ingreso en la Real Academia Española. Madrid. Cito por su reedición en la serie *Discursos de ingreso en la Real Academia Española*, Madrid: Biblioteca Nueva, 2014.
- Prieto García-Seco, David (2015), “*La pícara Justina* en el *Diccionario de Autoridades*”, in J. M. García Martín – F. J. de Cos Ruiz – M. Franco Figueroa (eds.): *Actas del IX Congreso internacional de historia de la lengua española (Cádiz, 2012)*. Madrid - Frankfurt: Iberoamericana – Vervuert, 1593 – 1614.
- Rojo, Guillermo (2014): «Análisis cuantitativo de las citas del *Diccionario de Autoridades*», BRAE, XCIV, pp. 137-196.
- Rojo, Guillermo (2016): «Análisis cuantitativo de las citas de obras en el *Diccionario de Autoridades*». En prensa.
- Ruhstaller, Stefan (2000): «Las autoridades del *Diccionario de Autoridades*», en Ruhstaller, Stefan y Josefina Prado Aragonés (eds.): *Tendencias en la investigación lexicográfica del español*, Universidad de Huelva, pp. 93-224.
- Salvador Rosa, Aurora (1985): «Las localizaciones geográficas en el *Diccionario de Autoridades*», LEA, 8, pp. 103-139.